

1
2
3
4
5
6
7 IN RE: UBER TECHNOLOGIES, INC.,
8 PASSENGER SEXUAL ASSAULT
9 LITIGATION

10 This Order Relates to:
11 ALL ACTIONS

12 Case No. 23-md-03084-CRB (LJC)

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28 **ORDER RESOLVING OUTSTANDING
ESI PROTOCOL DISPUTES**

Re: Dkt. 499

The Court previously resolved most of the parties' disputes as to their competing proposed protocols for electronically stored information (ESI) in Pretrial Order (PTO) No. 9. Dkt. 345. The Court provided instructions to the parties for further investigations and meet-and-confer regarding the cloud-stored documents issues, metadata fields, and related provisions of the ESI protocol. *Id.* at 26.¹ The parties now submit a Joint Discovery Letter regarding their remaining disagreements, which the Court addresses herein.² Dkt. 499.

Relevant to the current dispute concerning the ESI protocol in this case, Uber uses Google Workspace. *See* Decl. of William Anderson ¶ 3 (dkt. 499-4). Google Workspace provides Uber with a suite of cloud-based web applications, including Gmail, Google Chat, Google Drive, and Google Vault. Decl. of Jamie Brown ¶ 4 (dkt. 499-6). Uber utilizes Google Vault as an information governance and e-discovery tool for its Google Workspace data. Anderson Decl. ¶ 14. Through Google Vault, Uber retains and holds, among other things, users' Gmail messages and Google Drive files. *Id.* Google Vault can act as a clearinghouse to process data for discovery

¹ Unless specified otherwise, the Court refers to the PDF page number generated by the Court's e-filing system.

² In this Order, the Court refers to Defendants collectively as "Uber," as the Court and the parties have throughout this litigation.

1 purposes. Decl. of Philip Favro ¶ 18 (dkt. 262-8). Uber’s standard discovery process involves
2 exporting a custodian’s Gmail messages and Google Drive files from Google Vault. Anderson
3 Decl. ¶ 14. Data not stored using Google Vault is sometimes referred to as “active data”. *See*
4 Brown Decl. ¶¶ 26, 30; *see also* Joint Discovery Letter at 2 (Plaintiffs distinguishing between
5 “Active Google Email” versus “Google Vault Email,” and “Google Drive Documents” versus
6 “Google Vault Documents.”).

7 Of principle concern here, a Gmail or Google Chat message that contains a hyperlink to a
8 document is referencing a Google Drive document that may still be evolving. A recipient or
9 others may modify that referenced document, which is centrally located so multiple people can
10 access and edit it. Furthermore, Google Vault does not export, collect, or connect the
11 contemporaneous versions of hyperlinked documents with the corresponding emails or messages
12 in which they are found. Rather, when a hyperlinked Google Drive document is exported from
13 Google Vault, the current version of that document is exported. If a Google Drive document
14 archived using Google Vault was edited after the email with the hyperlink to the document was
15 sent, then the Google Vault export will not reflect the version of the document that existed at the
16 time of the email. For data archived using Google Vault, and no longer in the active Google
17 Workspace, there is a manual process in place to identify a historic version of a hyperlinked
18 Google Drive document contemporaneous with the email communication. Anderson Decl. ¶¶ 17–
19 45.³

20 Certain technologies have been developed to link email and chat messages to Google Drive
21 documents, but there are limitations. Metaspike’s Forensic Email Collector (FEC) program can
22 retrieve active Google Email and contemporaneous versions of linked Google Drive documents,
23 but it does not have the ability to do the same with Google Email and Drive documents archived
24 using Google Vault. Supplemental Decl. of Favro ¶ 13 (dkt. 499-5). Uber’s e-discovery vendor,
25 Lighthouse, has developed a tool, Google Parser, that extracts specific links to Google Drive
26 documents from email and chat messages and certain metadata. Brown Decl. ¶ 20. Google

27
28 ³ The steps in the manual process are described in extensive detail, but there is no information
concerning cost or length of time to conduct the process.

1 Parser facilitates the grouping together of a message and document stored in Google Drive for
 2 purposes of review and production, and it contains certain metadata fields relevant to search,
 3 review, and production of messages. *Id.* ¶¶ 22–23. However, there is no evidence that this
 4 technology, which is an extraction tool, has been refined and deployed to collect contemporaneous
 5 versions of hyperlinked documents archived with Google Vault.

6 In sum, the briefing and evidence, as well as related case law, have made clear that cloud
 7 computing and document retention through Google Drive and Google Vault introduce a host of
 8 challenges to producing hyperlinked documents from Google Drive and other sources. *See*
 9 *Nichols v. Noom Inc.*, No. 20CV3677LGSKHP, 2021 WL 948646, at *2 (S.D.N.Y. Mar. 11, 2021)
 10 (recognizing that “complex questions about what constitutes reasonable search and collection
 11 methods” result from “the changing nature of how documents are stored and should be
 12 collected.”). Yet, contemporaneous versions of hyperlinked documents can support an inference
 13 regarding “who knew what, when.” An email message with a hyperlinked document may reflect a
 14 logical single communication of information at a specific point in time, even if the hyperlinked
 15 document is later edited. Thus, important evidence bearing on claims and defenses may be at
 16 stake, but the ESI containing that evidence is not readily available for production in the same
 17 manner that traditional email attachments could be produced. *See* Favro Decl. ¶¶ 13–15
 18 (describing the challenges associated with collecting specific iterations of hyperlinked documents
 19 compared with traditional email attachments).

20 **I. CLOUD-STORED DOCUMENTS AND RELATED HYPERLINKS**

21 The parties’ proposed language for the ESI protocol as to cloud-stored documents is as
 22 follows:

23 Uber’s Proposal	24 Plaintiffs’ Proposal
25 Uber’s Proposed Section 17: Cloud Stored Documents	26 Plaintiffs’ Proposed Section 18(a)-(c): Cloud Stored Documents
27 Uber will make reasonable and proportionate efforts to preserve the metadata relationship between email messages with links to files on Google Drive, to the extent Uber’s vendor for processing and managing	28 a) <u>Metadata Preserved.</u> Uber shall preserve the metadata relationship between email messages with links to files on Google Drive. Uber shall preserve and produce (including, if necessary, as custom

1 the documents to be reviewed and produced in
 2 this action possesses technology that enables it
 3 to maintain such a relationship. Defendants
 4 may use Lighthouse's "Google Parser" for this
 5 purpose.

6 Notwithstanding that Uber agrees to make
 7 reasonable and proportionate efforts in this
 8 regard, because of technological limitations
 9 inherent in the processing of emails containing
 10 embedded links, it shall not be presumed that
 11 all emails containing links to files on Google
 12 Drive will be produced with a metadata
 13 relationship between the parent email and the
 14 linked document.

15 To the extent the Receiving Party believes
 16 that there is a lack of a metadata relationship
 17 between a specific email message and a
 18 specific linked document, the Receiving Party
 19 may notify the Producing Party and request
 20 that the particular linked file be extracted and
 21 produced or identified. To the extent that the
 22 linked file in question is nonprivileged, and is
 23 relevant to either Party's claims or defenses
 24 and the efforts required to search for it would
 25 be proportional to the needs of the case, the
 26 Producing Party shall use reasonable and
 27 proportionate efforts to collect and
 28 produce/identify the document that was linked
 in the original email. The Parties agree to meet
 and confer to resolve any disagreements as to
 what constitutes reasonable and proportional
 discovery efforts.

fields) all metadata collected with respect to
 all cloud-stored documents. That includes, but
 is not limited to, all metadata output by
 Google Vault when exporting a matter. Thus,
 the metadata exported from Google Vault
 pertaining to each document shall be preserved
 and produced as metadata for the same
 document within the load file of any
 production containing any such document.

b) Hyperlinked/URL-Referenced Documents.

Producing party shall make all reasonable
 efforts to maintain and preserve the
 relationship between any message or email
 and any cloud-hosted document hyperlinked or
 referenced within the message or email. Thus,
 for instance, where a collected email links to
 or references by URL a document on Google
 Drive (or housed within Google vault,) the
 metadata for that message or email shall
 include the URLs and Google Document ID of
 all hyperlinked documents, and if a
 hyperlinked document was produced.

**c) Contemporaneous Versions of
 Hyperlinked/URL-Referenced Documents.**

Uber shall produce the contemporaneous
 document version, i.e., the document version
 likely present at the time an email or message
 was sent, of Google Drive documents
 referenced by URL or hyperlinks in emails or
 messages such as Slack. If Uber contends that
 it is unable to meet this requirement through
 commercial or vendor software, Google APIs
 or through other reasonable manual or
 automated means, then Plaintiffs and
 Defendants shall meet and confer to discuss
 solutions. This will not exempt Uber from
 producing the applicable version of any
 document so referenced by URL or hyperlink.

23 In PTO No. 9, the Court ordered as follows:

24
 25 Uber shall direct an employee with knowledge and expertise
 26 regarding Google Vault and Uber's data and information systems to
 27 investigate in detail the extent to which Google Vault's API, macro
 28 readers, Metaspike's FEC or other programs may be useful to
 automate, to some extent, the process of collecting the
 contemporaneous version of the document linked to a Gmail or other
 communication within Uber's systems, whether the email or
 communication is stored in Google Vault, or outside. This

1 investigation shall not be limited to documents referenced by URL or
2 hyperlinks in emails or Google documents stored in Google Vault, but
3 shall also include other cloud-based messages such as Slack. Uber's
4 designated employee may consult with Uber's e-discovery experts.
Likewise, Plaintiffs shall also more thoroughly investigate these
5 potential solutions.

6 Dkt. 345 at 21. The Court ordered the parties to meet and confer as to their investigations, and if
7 disagreements remained, and they wanted to submit a discovery letter to resolve their dispute, then
8 the Uber employee designated to conduct its investigation was to "submit a declaration supporting
9 its positions, and Plaintiff's expert(s) and/or e-discovery vendors shall do the same." *Id.* Uber
10 was also permitted to "submit declarations from its experts and e-discovery vendors." *Id.* Uber
11 submitted declarations from William Anderson, an Uber employee who specializes in e-discovery;
12 Jamie Brown, a Lighthouse representative; and a supplemental declaration from Philip Favro,
13 Uber's e-discovery expert. Dkt. 499-4, 499-5, 499-6. Plaintiffs submitted a second declaration
14 from Douglas Forrest, their e-discovery expert. Dkt. 499-1.

15 Uber's position, based on its "exhaustive" investigation, is that "no technical, scalable
16 solution is available" to automate the process of collecting contemporaneous versions of
17 hyperlinked documents. Joint Discovery Letter at 4. Plaintiffs, on the other hand, claim to have
18 identified a process for automation. Plaintiffs submit a Proposed Methodology for Retrieving
19 Google Drive Documents Linked to Within Google Emails (Proposed Methodology). Dkt. 499-2.
20 They have created a Proof-of-Concept program, described in detail in Mr. Forrest's declaration,
21 which supposedly demonstrates that there is a method available to programmatically retrieve
22 contemporaneous versions of linked Google Drive documents. *See* Forrest Decl. ¶¶ 29–38. The
23 Proposed Methodology provides that Uber is to create such a program based on Plaintiffs' Proof
24 of Concept to produce contemporaneous versions of documents with Google Drive hyperlinks.
Dkt. 499-2 at 3.

25 The Court is not persuaded that Plaintiffs' Proposed Methodology is a reasonably available
26 option here. Mr. Forrest's consulting firm created the Proof-of-Concept program based on a post
27 in Stack Overflow, "a well-known and widely used forum for developers," which contains sample
28 scripts purportedly for retrieving "a date-specific revision of a Google Drive Google native

1 document identified by its Document ID.” Forrest Decl. ¶ 33. However, as Mr. Anderson points
2 out, the anonymous internet user who posted the script on which the Proof-of-Concept program is
3 based on admitted that it did not work. Anderson Decl. ¶ 62. And even “a functioning version of
4 the script would not address the issues presented here,” in part because the “script was designed
5 for a single document using the Google Drive API, restoring a non-Vault document, with owner
6 access. This script would not work for Google Vault.” *Id.* Uber and its experts also object to
7 many other aspects of the Proof-of-Concept program, which they claim “ignores necessary steps in
8 the collection process and disregards numerous points of manual intervention.” Joint Discovery
9 Letter at 5.

10 Perhaps there is a way to work out these kinks, as Plaintiffs suggest (*id.* at 3), and the
11 Proof-of-Concept program can eventually be used by some developer as a foundation for creating
12 a program to automate the process of collecting from archived Google Vault data the
13 contemporaneous versions of Google Drive hyperlinked documents. But the Court will not order
14 Uber to expend potentially significant time and resources to develop such a program in order to
15 produce discovery in this MDL, as the program’s effectiveness is not assured. *See Nichols*, 2021
16 WL 948646, at *4 (rejecting a similar proposal from Mr. Forrest that the defendant’s e-discovery
17 programmers could create a program to utilize Google’s API to “extract links to Google Drive
18 documents from other Google Drive documents, emails, and Slack communications,” in part
19 because Mr. Forrest’s declaration did not “address the time it would take to apply the program,
20 load, and review the documents.”). The Court is satisfied by Uber’s showing in the Joint
21 Discovery Letter that it has thoroughly investigated the issue, and that no technological solution is
22 currently readily available to automate the process when it comes to collecting contemporaneous
23 versions of hyperlinked documents.

24 The Court has reviewed Uber’s evidence regarding the manual process to query Google
25 Vault for contemporaneous versions of a hyperlinked document. *See e.g.*, Favro Supplemental
26 Decl. ¶¶ 6–7. The Court is mindful of the burdens to Uber but also recognizes that Uber has
27 chosen Google Vault as its storage method. *See Shenwick v. Twitter*, Case No. 4:16-cv-05314,
28 2018 WL 5735176, at *1 (N.D. Cal. Sept. 17, 2018) (requiring the defendants to conduct manual

1 searches and produce 200 hyperlinked documents despite the undisputed burdens because they
2 chose their storage method and the plaintiffs are entitled to discovery). In *Shenwick*, the
3 defendants utilized Google Vault for document review and production. *See id.*, Dkt. 190 at 5.
4 Thus, the potential limitations and pitfalls with respect to production of hyperlinked documents
5 from Google Vault have been widely known for many years, yet Uber has elected to transfer and
6 retain its electronic data using this service.

7 Given all the above, the ESI protocol shall state the following with respect to Cloud Stored
8 Documents:

9 **a) Metadata Preserved.**

10 Uber shall preserve the metadata relationship between email
11 messages with links to files on Google Drive to the extent feasible
12 with existing technology. Uber shall preserve and produce (including,
13 if necessary, as custom fields) all metadata collected with respect to
14 all cloud-stored documents. That includes, but is not limited to, all
15 metadata output by Google Vault when exporting a matter. Thus, the
16 metadata exported from Google Vault pertaining to each document
17 shall be preserved and produced as metadata for the same document
18 within the load file of any production containing any such document.

19 **b) Hyperlinked/URL-Referenced Documents.**

20 Producing party shall make all reasonable efforts to maintain and
21 preserve the relationship between any message or email and any
22 cloud-hosted document hyperlinked or referenced within the message
23 or email. Thus, for instance, where a collected email links to or
24 references by URL a document on Google Drive (or housed within
25 Google vault,) the metadata for that message or email shall include
26 the URLs and Google Document ID of all hyperlinked documents.

27 **c) Contemporaneous Versions of Hyperlinked/URL-Referenced**
Documents.

28 Uber shall produce, to the extent feasible on an automated, scalable
basis with existing technology, the contemporaneous document
version, i.e., the document version likely present at the time an email
or message was sent, of Google Drive documents referenced by URL
or hyperlinks therein. For hyperlinked Google Workspace data
archived using Google Vault, Uber is not required to produce the
contemporaneous document version at the time the email or message
was sent, as this is not possible through an automated process with
existing technology. However, Plaintiffs may identify up to 200
hyperlinks for which they seek the contemporaneous referenced
document even though the email or message has been archived with
Google Vault. Uber shall identify and produce the likely
contemporaneous versions that Plaintiffs have requested. The scope
of this production does not exempt Uber from any obligation that it
preserve historic versions or revision history of any document
referenced by URL or hyperlink.

1 The parties may seek to modify the protocol with respect to the production of
2 contemporaneous versions of hyperlinked documents based on the need for relevant discovery and
3 what is proportional to the needs of the case, or what electronic data is not reasonably accessible
4 because of undue burden or cost.⁴ By stipulation of the parties, or through Court order, Plaintiffs
5 may request additional contemporaneous documents, and Uber may seek relief from the
6 production of certain versions or other obligations under the ESI protocol based on undue burden
7 or costs, overbreadth or disproportionality.

8 **II. METADATA FIELDS**

9 PTO No. 9 required the parties to meet and confer as to certain disputed metadata fields,
10 but proposed a resolution of the disputes if they were unable to come to an agreement. Dkt. 345 at
11 25. The parties now indicate that they reached an agreement as to three of the disputed metadata
12 fields. Joint Discovery Letter at 6. The fields still in dispute are LINKGOOGLEDRIVEURLS
13 and the Account field. *Id.* In PTO No. 9, the Court proposed that the
14 LINKGOOGLEDRIVEURLS and Account fields be excluded. Dkt. 345 at 25.

15 The Court first addresses the disputed LINKGOOGLEDRIVEURLS metadata field. As
16 Plaintiffs' expert Mr. Forrest explains it, this is a custom created metadata field designed to
17 identify which Google Drive documents, if any for whatever reason, have not been produced.
18 Forrest Decl. ¶ 95. Uber contends that Plaintiffs did not promptly follow up in regard to this
19 metadata field, and only made related inquiries the day before the parties' joint discovery letter
20 was due. Joint Discovery Letter at 6. Uber also responds that the custom creation of the
21 LINKGOOGLEDRIVEURLS would impose a substantial burden. *Id.*

22 Plaintiffs now propose an alternate approach instead of creating the
23 LINKGOOGLEDRIVEURLS metadata field. The proposed alternate approach involves adding a

24
25 ⁴ Plaintiffs provide data estimates to flag the large volume of hyperlinked documents that will be
26 produced in non-contemporaneous form. Joint Discovery Letter at 2. The Court is unable to
27 assess how meaningful these estimates are, given that there may be duplication among the
28 hyperlinked documents, and that multiple emails may refer to the same document that is rarely
edited over a period of time. The Court has concerns regarding the difficulty of securing the
version of a document that is contemporaneous with the email that hyperlinked it. However, the
full scope of the problem is not yet defined, and typically only a fraction of documents produced
in discovery are material to the litigation.

1 metadata field, “Missing Google Drive Attachments,” and a second metadata field, “Non-
2 Contemporaneous.” Forrest Decl. ¶ 96. Uber did not present any evidence indicating that the
3 alternate proposal would impose a substantial burden. Nor did Uber state that it was unable to
4 respond to the alternate proposal because Plaintiffs did not raise it in sufficient time. The first new
5 metadata field proposed by Plaintiffs, “Missing Google Drive Attachments,” identifies the Google
6 Drive documents that were not produced by providing links to all linked Google Drive documents
7 that could not be retrieved. *Id.* The second new metadata field proposed by Plaintiffs identifies
8 the produced Google Drive documents that were not the versions in existence at the time the
9 document was hyperlinked, i.e., the “non-contemporaneous” version of the document that was
10 produced. *Id.*

11 As noted previously, the Court directed the parties to meet and confer regarding disputed
12 metadata fields, which included the LINKGOOGLEDRIIVEURLS field. The Court did not intend
13 to broaden negotiations regarding what metadata fields should be included in the ESI protocol.
14 The Court recognizes, however, that early ESI decisions may need to be modified when new
15 information is learned. “[P]roportionality considerations may need to be re-balanced at later
16 points in the litigation, and...discovery plans may be modified when new information is learned.”
17 *Nichols*, 2021 WL 948646 at *3. In this Order, the Court has declined to require Uber to produce
18 contemporaneous versions of every hyperlinked document.

19 Plaintiffs’ new metadata proposals appear helpful to streamline review of Uber’s
20 productions because the metadata will identify which hyperlinked documents are missing from the
21 production and which documents produced are the non-contemporaneous versions. Streamlining
22 review of the production of hyperlinked documents will advance the speedy and less expensive
23 determination of this action. Fed. R. Civ. P. 1. Furthermore, Uber, as a responding party, is
24 bound by its obligations under Rule 26(g) of the Federal Rules of Civil Procedure. An attorney’s
25 signature pursuant to Rule 26(g) “certifies that the lawyer has made a reasonable effort to assure
26 that the client has provided all of the information and documents available to him that are
27 responsive to the discovery demand.” Fed. R. Civ. P. 26, Notes of Advisory Committee on
28 Rules—1983 Amendment. Conversely, parties should not demand metadata fields for which they

1 have no practical use, or which do not materially aid in the discovery process. *See The Sedona*
 2 *Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing*
 3 *Electronic Document Production*, 19 Sedona Conf. J. 1, 173 (2018). “[T]he parties should follow

4 the Federal Rules and seek to reach agreement for the production of ESI as ‘ordinarily maintained

5 or in a reasonably usable form or forms’ that provide the requesting party a functionally adequate

6 ability to access, cull, analyze, search, and display the ESI, as may be appropriate given its nature

7 and the proportional needs of the case.” *Id.* at 175. Given the challenges associated with the

8 production of hyperlinked documents from Google Vault, Plaintiffs’ request to add the metadata

9 field “Missing Google Drive Attachments,” and a second metadata field, “Non-

10 Contemporaneous,” is granted.

11 Plaintiffs’ request for the Account field without the limitation proposed by Uber (Forrest

12 Decl. ¶ 92) is denied, as Plaintiffs have not explained why knowing the particular email address

13 associated with a custodian’s responsive documents will materially advance the development of

14 their case. Plaintiffs will already be informed of the custodian’s first and last name, and their

15 primary email address, and may inquire through other discovery methods whether other email

16 addresses were used. Plaintiffs’ request is denied without prejudice to a future, more limited

17 metadata request where Plaintiffs demonstrate that such granular data is material to the discovery

18 process.

19 **III. DEFINITIONS**

20 The parties were unable to resolve their dispute over the definition of “attachment.” Dkt.

21 499 at 7. The proposed language for the ESI protocol is as follows:

Uber’s Proposal	Plaintiffs’ Proposal
Uber’s Proposed Definition of “Attachment”	Plaintiffs’ Proposed Definition of “Attachment”
An “Attachment” is typically a file associated with another file for the purpose of storage, transfer, processing, production, or review. There may be multiple attachments associated with a single “parent” or “master” file. In many records and information management	“Attachment(s)” shall be interpreted broadly and includes, e.g., traditional email attachments and documents embedded in other documents (e.g., Excel files embedded in PowerPoint files) as well as modern attachments, pointers, internal or non-public

1	systems, or in a litigation context, the 2 attachments and associated record(s) may be 3 managed and processed as a single unit. In 4 common use, this term most often refers to a 5 file (or files) associated with an individual 6 email or other message type. For the 7 avoidance of doubt, a hyperlinked document, 8 such as a cloud-based document in Google 9 Drive, is not an “attachment.”	10 documents linked, hyperlinked, stubbed or 11 otherwise pointed to within or as part of other 12 ESI (including but not limited to email, 13 messages, comments or posts, or other 14 documents).
---	---	---

7 Plaintiffs' proposal is adopted with the following modification:

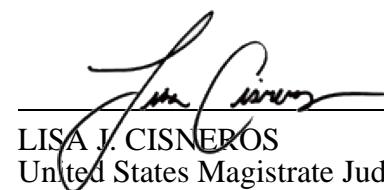
8 “Attachment(s)” shall be interpreted broadly and includes, e.g.,
9 traditional email attachments and documents embedded in other
10 documents (e.g., Excel files embedded in PowerPoint files) as well as
11 modern attachments, pointers, internal or non-public documents
12 linked, hyperlinked, stubbed or otherwise pointed to within or as part
13 of other ESI (including but not limited to email, messages, comments
14 or posts, or other documents). This definition does not obligate Uber
15 to produce the contemporaneous version of Google Drive documents
16 referenced by URL or hyperlinks if no existing technology makes it
17 feasible to do so.

IV. CONCLUSION

Having resolved all the remaining disputes as to the terms of the ESI protocol, the parties are hereby ordered to incorporate the Court's rulings from PTO No. 9 and from this Order into a final Stipulated and [Proposed] ESI Protocol, to be submitted to the Court by April 30, 2024.

IT IS SO ORDERED.

Dated: April 23, 2024



LISA J. CISNEROS
United States Magistrate Judge